

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: SEED, BRIAN
HAAS, JURGEN
- (ii) TITLE OF THE INVENTION: HIGH LEVEL EXPRESSION OF
PROTEINS
- (iii) NUMBER OF SEQUENCES: 110
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Clark & Elbing LLP
 - (B) STREET: 176 Federal Street
 - (C) CITY: Boston
 - (D) STATE: MA
 - (E) COUNTRY: USA
 - (F) ZIP: 02110
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/717,294
 - (B) FILING DATE: 20-SEP-1996
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Elbing, Karen L
 - (B) REGISTRATION NUMBER: 35,238
 - (C) REFERENCE/DOCKET NUMBER: 00786/345001
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 617-428-0200
 - (B) TELEFAX: 617-428-7045
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Other
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CGCGGGCTAG CCACCGAGAA GCTG

24

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| | |
|--|-----|
| ACCGAGAAGC TGTGGGTGAC CGTGTACTAC GCGGTGCCCCG TGTGGAAGAG AGGCCACCAC | 60 |
| CACCCTGTTT TGCGCCAGCG ACGCCAAGGC GTACGACACC GAGGTGCACA ACGTGTGGGC | 120 |
| CACCCAGGCG TCGGTGCCCA CCGACCCCAA CCCCAGGAG GTGGAGCTCG TGAACGTGAC | 180 |
| CGAGAACTTC AACAT | 195 |

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| | |
|---------------------------------------|----|
| CCACCATGTT GTTCTTCCAC ATGTTGAAGT TCTC | 34 |
|---------------------------------------|----|

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

| | |
|--------------------------------------|----|
| GACCGAGAAC TTCAACATGT GGAAGAACAA CAT | 33 |
|--------------------------------------|----|

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

| | |
|---|-----|
| TGGAAGAACA ACATGGTGGA GCAGATGCAT GAGGACATCA TCAGCCTGTG GGACCAGAGC | 60 |
| CTGAAGCCCT GCGTGAAGCT GACCCCTGT GCGTGACCTG AACTGCACCG ACCTGAGGAA | 120 |
| CACCAACAA ACACACAG CACCGCCAAC AACACAGCA ACAGCGAGGG CACCATCAAG | 180 |
| GGCGGCGAGA TG | 192 |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTTGAAGCTG CAGTTCTTCA TCTCGCCGCC CTT

33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 31 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GAAGAACTGC AGCTTCAACA TCACCACCAG C

31

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 195 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

| | |
|---|-----|
| AACATCACCA CCAGCATCCG CGACAAGATG CAGAAGGAGT ACGCCCTGCT GTACAAGCTG | 60 |
| GATATCGTGA GCATCGACAA CGACAGCACC AGCTACCGCC TGATCTCCTG CAACACCAGC | 120 |
| GTGATCACCC AGGCCTGCCC CAAGATCAGC TTCGAGCCCA TCCCCATCCA CTACTGCGCC | 180 |
| CCCGCCGGCT TCGCC | 195 |

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAACTTCTTG TCGGCGGCGA AGCCGGCGGG

30

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GCGCCCCCGC CGGCTTCGCC ATCCTGAAGT GCAACGACAA GAAGTTC

47

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

| | |
|---|-----|
| GCCGACAAGA AGTTCAGCGG CAAGGGCAGC TGCAAGAACG TGAGCACCGT GCAGTGCACC | 60 |
| CACGGCATCC GGCCGGTGGT GAGCACCCAG CTCCTGCTGA ACGGCAGCCT GGCCGAGGAG | 120 |
| GAGGTGGTGA TCCGCAGCGA GAACTTCACC GACAACGCCA AGACCATCAT CGTGCACCTG | 180 |
| AATGAGAGCG TGCAGATC | 198 |

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 34 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AGTTGGGACG CGTGCAGTTG ATCTGCACGC TCTC 34

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGAGCGTGC AGATCAACTG CACGCGTCCC 30

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 120 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

| | |
|---|-----|
| AACTGCACGC GTCCCAACTA CAACAAGCGC AAGCGCATCC ACATCGGCCC CGGGCGCGCC | 60 |
| TTCTACACCA CCAAGAACAT CATCGGCACC ATCCTCCAGG CCCACTGCAA CATCTCTAGA | 120 |

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GTCGTTCCAC TTGGCTCTAG AGATGTTGCA

30

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCAACATCTC TAGAGCCAAG TGGAACGAC

29

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 131 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GCCAAGTGGA ACGACACCCT GCGCCAGATC GTGAGCAAGC TGAAGGAGCA GTTCAAGAAC
AAGACCATCG TGTTCAACCAG AGCAGCGGCG GCGACCCCGA GATCGTGATG CACAGCTTCA
ACTGCGGCGG C

60
120
131

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCAGTAGAAG AATTCGCCGC CGCAGTTGA

29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCAACTGCGG CGGCGAATTC TTCTACTGC

29

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

| | |
|---|-----|
| GGCGAATTCT TCTACTGCAA CACCAGCCCC CTGTTCAACA GCACCTGGAA CGGCAACAAC | 60 |
| ACCTGGAACA ACACCACCGG CAGCAACAAC AATATTACCC TCCAGTGCAA GATCAAGCAG | 120 |
| ATCATCAACA TGTGGCAGGA GGTGGGCAAG GCCATGTACG CCCCCCCAT CGAGGGCCAG | 180 |
| ATCCGGTGCA GCAGC | 195 |

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GCAGACCGGT GATGTTGCTG CTGCACCGGA TCTGGCCCTC

40

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CGAGGGCCAG ATCCGGTGCA GCAGCAACAT CACCGGTCTG

40

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 198 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

| | |
|--|-----|
| AACATCACCG GTCTGCTGCT GACCCGCGAC GGCGGCAAGG ACACCGACAC CAACGACACC | 60 |
| GAAATCTTCC GCCCCGGCGG CGGCGACATG CGCGACAACCT GGAGATCTGA GCTGTACAAG | 120 |
| TACAAGGTGG TGACGATCGA GCCCCTGGGC GTGGCCCCCA CCAAGGCCAA GCGCCGCGTG | 180 |
| GTGCAGCGCG AGAAGCGC | 198 |

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGCGGGCGGC CGCTTTAGCG CTTCTCGCGC TGCACCAC 38

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCGGGGGAT CCAAGCTTAC CATGATTCCA GTAATAAGT 39

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 165 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

ATGAATCCAG TAATAAGTAT AACATTATTA TTAAGTGTAT TACAAATGAG TAGAGGACAA 60
AGAGTAATAA GTTTAACAGC ATCTTTAGTA AATCAAAATT TGAGATTAGA TTGTAGACAT 120
GAAAATAATA CAAATTTGCC AATACAACAT GAATTTTCAT TAACG 165

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CGCGGGGAAT TCACGCGTTA ATGAAAATTC ATGTTG 36

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CGCGGATCCA CGCGTGAAAA AAAAAACAT

30

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 150 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CGTGAAAAAA AAAAAACATGT ATTAAGTGGA ACATTAGGAG TACCAGAACA TACATATAGA
AGTAGAGTAA ATTTGTTTAG TGATAGATTC ATAAAAGTAT TAACATTAGC AAATTTTACA
ACAAAAGATG AAGGAGATTA TATGTGTGAG

60

120

150

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CGCGAATTCG AGCTCACACA TATAATCTCC

30

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CGCGGATCCG AGCTCAGAGT AAGTGGACAA

30

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 170 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTCAGAGTAA GTGGACAAAA TCCAACAAGT AGTAATAAAA CAATAAATGT AATAAGAGAT

60

| | | | | | | |
|-------------|------------|------------|------------|------------|------------|-----|
| AAATTAGTAA | AATGTGAGGA | ATAAGTTTAT | TAGTACAAAA | TACAAGTTGG | TTATTATTAT | 120 |
| TATTATTAAAG | TTTAAGTTTT | TTACAAGCAA | CAGATTTTAT | AAGTTTATGA | | 170 |

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

| | | | | |
|------------|------------|------------|--------|----|
| CGCGAATTCG | CGGCCGCTTC | ATAAACTTAT | AAAATC | 36 |
|------------|------------|------------|--------|----|

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1632 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

| | | | | | | |
|------------|------------|------------|------------|-------------|------------|------|
| CTCGAGATCC | ATTGTGCTCT | AAAGGAGATA | CCCGGCCAGA | CACCCTCACC | TGCGGTGCCC | 60 |
| AGCTGCCCAG | GCTGAGGCAA | GAGAAGGCCA | GAAACCATGC | CCATGGGGTC | TCTGCAACCG | 120 |
| CTGGCCACCT | TGTACCTGCT | GGGGATGCTG | GTCGCTTCCG | TGCTAGCCAC | CGAGAAGCTG | 180 |
| TGGGTGACCG | TGTACTACGG | CGTGCCCGTG | TGGAAGGAGG | CCACCACCAC | CCTGTTCTGC | 240 |
| GCCAGCGACG | CCAAGGCGTA | CGACACCGAG | GTGCACAACG | TGTGGGGCCAC | CCAGGCGTGC | 300 |
| GTGCCCACCG | ACCCCAACCC | CCAGGAGGTG | GAGCTCGTGA | ACGTGACCGA | GAACTTCAAC | 360 |
| ATGTGGAAGA | ACAACATGGT | GGAGCAGATG | CATGAGGACA | TCATCAGCCT | GTGGGACCAG | 420 |
| AGCCTGAAGC | CCTGCGTGAA | GCTGACCCCC | CTGTGCGTGA | CCCTGAACTG | CACCGACCTG | 480 |
| AGGAACACCA | CCAACACCAA | CAACAGCACC | GCCAACAACA | ACAGCAACAG | CGAGGGCACC | 540 |
| ATCAAGGGCG | GCGAGATGAA | CAACTGCAGC | TTCAACATCA | CCACCAGCAT | CCGCGACAAG | 600 |
| ATGCAGAAGG | AGTACGCCCT | GCTGTACAAG | CTGGATATCG | TGAGCATCGA | CAACGACAGC | 660 |
| ACCAGCTACC | GCCTGATCTC | CTGCAACACC | AGCGTGATCA | CCCAGGCCTG | GCCCAAGATC | 720 |
| AGCTTCGAGC | CCATCCCCAT | CCACTACTGC | GCCCCGCCG | GCTTCGCCAT | CCTGAAGTGC | 780 |
| AACGACAAGA | AGTTCAGCGG | CAAGGGCAGC | TGCAAGAACG | TGAGCACCGT | GCAGTGCACC | 840 |
| CACGGCATCC | GGCCGGTGGT | GAGCACCCAG | CTCCTGCTGA | ACGGCAGCCT | GGCCGAGGAG | 900 |
| GAGGTGGTGA | TCCGCAGCGA | GAACTTCACC | GACAACGCCA | AGACCATCAT | CGTGCACCTG | 960 |
| AATGAGAGCG | TGCAGATCAA | CTGCACGCGT | CCCAACTACA | ACAAGCGCAA | GCGCATCCAC | 1020 |
| ATCGGCCCCG | GGCGCGCCTT | CTACACCACC | AAGAACATCA | TCGGCACCAT | CCGCCAGGCC | 1080 |
| CACTGCAACA | TCTCTAGAGC | CAAGTGGAAC | GACACCCTGC | GCCAGATCGT | GAGCAAGCTG | 1140 |
| AAGGAGCAGT | TCAAGAACAA | GACCATCGTG | TTCAACCAGA | GCAGCGGCGG | CGACCCCGAG | 1200 |
| ATCGTGATGC | ACAGCTTCAA | CTGCGGCGGC | GAATTCTTCT | ACTGCAACAC | CAGCCCCCTG | 1260 |
| TTCAACAGCA | CCTGGAACGG | CAACAACACC | TGGAACAACA | CCACCGGCAG | CAACAACAAT | 1320 |
| ATTACCCTCC | AGTGCAAGAT | CAAGCAGATC | ATCAACATGT | GGCAGGAGGT | GGGCAAGGCC | 1380 |
| ATGTACGCCC | CCCCATCGA | GGGCCAGATC | CGGTGCAGCA | GCAACATCAC | CGGTCTGCTG | 1440 |
| CTGACCCGCG | ACGGCGGCAA | GGACACCGAC | ACCAACGACA | CCGAAATCTT | CCGCCCCGGC | 1500 |
| GGCGGCGACA | TGCGCGACAA | CTGGAGATCT | GAGCTGTACA | AGTACAAGGT | GGTGACGATC | 1560 |
| GAGCCCCTGG | GCGTGGCCCC | CACCAAGGCC | AAGCGCCGCG | TGGTGCAGCG | CGAGAAGCGC | 1620 |
| TAAAGCGGCC | GC | | | | | 1632 |

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2481 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|------|
| ACCGAGAAGC | TGTGGGTGAC | CGTGACTAC | GGCGTGCCCG | TGTGGAAGGA | GGCCACCACC | 60 |
| ACCCTGTTCT | GCGCCAGCGA | CGCCAAGGCG | TACGACACCG | AGGTGCACAA | CGTGTGGGCC | 120 |
| ACCCAGGCGT | GCGTGCCAC | CGACCCCAAC | CCCCAGGAGG | TGGAGCTCGT | GAACGTGACC | 180 |
| GAGAACTTCA | ACATGTGGAA | GAACAACATG | CTGGAGCAGA | TGCATGAGGA | CATCATCAGC | 240 |
| CTGTGGGACC | AGAGCCTGAA | GCCCTGCGTG | AAGCTGACCC | CCCTGTGCGT | GACCCCTGAAC | 300 |
| TGACCGGACC | TGAGGAACAC | CACCAACACC | AACAACAGCA | CCGCCAACAA | CAACAGCAAC | 360 |
| AGCGAGGGCA | CCATCAAGGG | CGGCGAGATG | AAGAACTGCA | GCTTCAACAT | CACCACCAGC | 420 |
| ATCCGCGACA | AGATGCAGAA | GGAGTACGCC | CTGCTGTACA | AGCTGGATAT | CGTGAGCATC | 480 |
| CACAACGACA | GCACCAGCTA | CCGCTGATC | TCCTGCAACA | CCAGCGTGAT | CACCCAGGCC | 540 |
| TGCCCCAAGA | TCAGCTTCGA | GCCCATCCCC | ATCCACTACT | GCGCCCCCGC | CGGCTTCGCC | 600 |
| ATCCTGAAGT | GCAACGACAA | GAAGTTCAGC | GGCAAGGGCA | GCTGCAAGAA | CGTGACCACC | 660 |
| GTGCAGTGCA | CCCACGGCAT | CCGGCCGGTG | GTGAGCACCC | AGCTCCTGCT | GAACGGCAGC | 720 |
| CTGGCCGAGG | AGGAGGTGGT | GATCCGCAGC | GAGAACTTCA | CCGACAACGC | CAAGACCATC | 780 |
| ATCGTGACCC | TGAATGAGAG | CGTGCAGATC | AAGTGCACGC | GTCCCAACTA | CAACAAGCGC | 840 |
| AAGCGCATCC | ACATCGGCCC | CGGGCGCGCC | TTCTACACCA | CCAAGAACAT | CATCGGCACC | 900 |
| ATCCGCCAGG | CCCACTGCAA | CATCTCTAGA | GCCAAGTGGA | ACGACAACCT | GCGCCAGATC | 960 |
| GTGAGCAAGC | TGAAGGAGCA | GTTCAAGAAC | AAGACCATCG | TGTTCAACCA | GAGCAGCGGC | 1020 |
| GGCGACCCCG | AGATCGTGAT | GCACAGCTTC | AACTGCGGCG | GCGAATTCTT | CTACTGCAAC | 1080 |
| ACCAGCCCCC | TGTTCAACAG | CACCTGGAAC | GGCAACAACA | CCTGGAACAA | CACCACCGGC | 1140 |
| AGCAACAACA | ATATTACCCT | CCAGTGCAAG | ATCAAGCAGA | TCATCAACAT | GTGGCAGGAG | 1200 |
| GTGGGCAAGG | CCATGTACGC | CCCCCCCATC | GAGGGCCAGA | TCCGGTGCA | GAGCAACATC | 1260 |
| ACCGGTCTGC | TGCTGACCCG | CGACGGCGGC | AAGGACACCG | ACACCAACGA | CACCGAAATC | 1320 |
| TTCCGCCCCG | GCGGCGGCGA | CATGCGCGAC | AACTGGAGAT | CTGAGCTGTA | CAAGTACAAG | 1380 |
| GTGGTGACGA | TCGAGCCCCC | GGGCGTGGCC | CCACCAAGG | CCAAGCGCCG | CGTGGTGCA | 1440 |
| CGCGAGAAGC | GGGCGGCCAT | CGGCGCCCTG | TTCTGGGGCT | TCCTGGGGGC | GGCGGGCAGC | 1500 |
| ACCATGGGGG | CGCCAGCGT | GACCCTGACC | GTGCAGGCCC | GCCTGCTCCT | GAGCGGCATC | 1560 |
| GTGCAGCAGC | AGAACAACCT | CCTCCGCGCC | ATCGAGGCCC | AGCAGCATAT | GCTCCAGCTC | 1620 |
| ACCGTGTGGG | GCATCAAGCA | GCTCCAGGCC | CGCGTGCTGG | CCGTGGAGCG | CTACCTGAAG | 1680 |
| GACCAGCAGC | TCCTGGGCTT | CTGGGGCTGC | TCCGGCAAGC | TGATCTGCAC | CACCACGGTA | 1740 |
| CCCTGGAACG | CCCTCTGGAG | CAACAAGAGC | CTGGACGACA | TCTGGAACAA | CATGACCTGG | 1800 |
| ATGCAGTGGG | AGCGCGAGAT | CGATAACTAC | ACCAGCCTGA | TCTACAGCCT | GCTGGAGAAG | 1860 |
| AGCCAGACCC | AGCAGGAGAA | GAACGAGCAG | GAGCTGCTGG | AGCTGGACAA | CTGGGCGAGC | 1920 |
| CTGTGGAAC | GGTTCGACAT | CACCAACTGG | CTGTGGTACA | TCAAAATCTT | CATCATGATT | 1980 |
| GTGGGCGGCC | TGGTGGGCCT | CCGCATCGTG | TTCCGCCGTG | TGAGCATCGT | GAACCGCGTG | 2040 |
| CGCCAGGGGT | ACAGCCCCCT | GAGCCTCCAG | ACCCGGCCCC | CCGTGCCCGC | CGGGCCCGAC | 2100 |
| CGCCCCGAGG | GCATCGAGGA | GGAGGGCGGC | GAGCGCGACC | GCGACACCAG | CGGCAGGCTC | 2160 |
| GTGCACGGCT | TCCTGGCGAT | CATCTGGGTC | GACCTCCGCA | GCCTGTTCTT | GTTTCAGCTAC | 2220 |
| CACCACCGCG | ACCTGCTGCT | GATCGCCGCC | CGCATCGTGG | AACTCCTAGG | CCGCCCGCGC | 2280 |
| TGGGAGGTGC | TGAAGTACTG | GTGGAACCTC | CTCCAGTATT | GGAGCCAGGA | GCTGAAGTCC | 2340 |
| AGCGCCGTGA | GCCTGCTGAA | CGCCACCGCC | ATCGCCGTGG | CCGAGGGCAC | CGACCGCGTG | 2400 |
| ATCGAGGTGC | TCCAGAGGGC | CGGGAGGGCG | ATCCTGCACA | TCCCCACCCG | CATCCGCCAG | 2460 |
| GGGCTCGAGA | GGGCGCTGCT | G | | | | 2481 |

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 486 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| ATGAATCCAG | TAATAAGTAT | AACATTATTA | TTAAGTGTAT | TACAAATGAG | TAGAGGACAA | 60 |
| AGAGTAATAA | GTTTAACAGC | ATGTTTAGTA | AATCAAAATT | TGAGATTAGA | TTGTAGACAT | 120 |
| GAAAATAATA | CACCTTTGCC | AATACAACAT | GAATTTTCAT | TAACGCGTGA | AAAAAAAAAA | 180 |
| CATGTATTAA | GTGGAACATT | AGGAGTACCA | GAACATACAT | ATAGAAGTAG | AGTAAATTTG | 240 |
| TTTAGTGATA | GATTCATAAA | AGTATTAACA | TTAGCAAAAT | TTACAACAAA | AGATGAAGGA | 300 |

| | | | | | | |
|------------|------------|-------------|------------|------------|------------|-----|
| GATTATATGT | GTGAGCTCAG | AGTAAGTGGA | CAAAATCCAA | CAAGTAGTAA | TAAAACAATA | 360 |
| AATGTAATAA | GAGATAAATT | AGTAAAAATGT | GGAGGAATAA | GTTTATTAGT | ACAAAATACA | 420 |
| AGTTGGTTAT | TATTATTATT | ATTAAGTTTA | AGTTTTTTAC | AAGCAACAGA | TTTTATAAGT | 480 |
| TTATGA | | | | | | 486 |

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

| | | | | | | |
|------------|------------|------------|------------|------------|------------|-----|
| ATGAACCCAG | TCATCAGCAT | CACTCTCCTG | CTTTCAGTCT | TGCAGATGTC | CCGAGGACAG | 60 |
| AGGGTGATCA | GCCTGACAGC | CTGCCTGGTG | AACAGAACCT | TCGACTGGAC | TGCCGTCATG | 120 |
| AGAATAACAC | CAACTTGCCC | ATCCAGCATG | AGTTCAGCCT | GACCCGAGAG | AAGAAGAAGC | 180 |
| ACGTGCTGTC | AGGCACCCTG | GGGGTTCCCG | AGCACACTTA | CCGCTCCCGC | GTCAACCTTT | 240 |
| TCAGTGACCG | CTTTATCAAG | GTCTTACTC | TAGCCAACTT | GACCACCAAG | GATGAGGGCG | 300 |
| ACTACATGTG | TGAACTTCGA | GTCTCGGGCC | AGAATCCCAC | AAGCTCCAAT | AAAACTATCA | 360 |
| ATGTGATCAG | AGACAAGCTG | GTCAAGTGTG | GTGGCATAAG | CCTGCTGGTT | CAAAACACTT | 420 |
| CCTGGCTGCT | GCTGCTCCTG | CTTCCCTCT | CCTTCCTCCA | AGCCACGGAC | TTCATTTCTC | 480 |
| TGTGA | | | | | | 485 |

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

| | | | | |
|------------|------------|------------|-----|----|
| CGCGGGGCTA | GCGCAAAGAG | TAATAAGTTT | AAC | 33 |
|------------|------------|------------|-----|----|

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

| | | | |
|------------|------------|----------|----|
| CGCGGATCCC | TTGTATTTTG | TACTAATA | 28 |
|------------|------------|----------|----|

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 762 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

| | | | | | | |
|------------|------------|------------|------------|------------|-------------|-----|
| GAATTCACGC | GTAAGCTTGC | CGCCACCATG | GTGAGCAAGG | GCGAGGAGCT | GTTCAACGGG | 60 |
| GTGGTGCCCA | TCCTGGTCGA | GCTGGACGGC | GACGTGAACG | GCCACAAGTT | CAGCGTGTCC | 120 |
| GGCGAGGGCG | AGGGCGATGC | CACCTACGGC | AAGCTGACCC | TGAAGTTCAT | CTGCACCACC | 180 |
| GGCAAGCTGC | CCGTGCCCTG | GCCCACCCTC | GTGACCACCT | TCAGCTACGG | CGTGCACTGC | 240 |
| TTCAGCCGCT | ACCCCGACCA | CATGAAGCAG | CACGACTTCT | TCAAGTCCGC | CATGCCCCGAA | 300 |
| GGCTACGTCC | AGGAGCGCAC | CATCTTCTTC | AAGGACGACG | GCAACTACAA | GACCCGCGCC | 360 |
| GAGGTGAAGT | TCGAGGGCGA | CACCCTGGTG | AACCGCATCG | AGCTGAAGGG | CATCGACTTC | 420 |
| AAGGAGGACG | GCAACATCCT | GGGGCACAAG | CTGGAGTACA | ACTACAACAG | CCACAACGTC | 480 |
| TATATCATGG | CCGACAAGCA | GAAGAACGGC | ATCAAGGTGA | ACTTCAAGAT | CCGCCACAAC | 540 |
| ATCGAGGACG | GCAGCGTGCA | GCTCGCCGAC | CACTACCAGC | AGAACACCCC | CATCGGCGAC | 600 |
| GGCCCCGTGC | TGCTGCCCCG | CAACCACTAC | CTGAGCACCC | AGTCCGCCCT | GAGCAAAGAC | 660 |
| CCCAACGAGA | AGCGCGATCA | CATGGTCCTG | CTGGAGTTCT | TGACCGCCGC | CGGGATCACT | 720 |
| CACGGCATGG | ACGAGCTGTA | CAAGTAAAGC | GGCCGCGGAT | CC | | 762 |

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4670 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

| | | | | | | |
|-------------|------------|------------|-------------|------------|-------------|------|
| AAGCTTAAAC | CATGCCCATG | GGGTCTCTGC | AACCGCTGGC | CACCTTGTAC | CTGCTGGGGA | 60 |
| TGCTGGTCGC | TTCCGTGCTA | GCCGCCACCA | GAAGATACTA | CCTGGGTGCA | GTGGAACTGT | 120 |
| CATGGGACTA | TATGCAAAGT | GATCTCGGTG | AGCTGCCTGT | GGACGCAAGA | TTTCCTCCTA | 180 |
| GAGTGCCAAA | ATCTTTTCCA | TTCAACACCT | CAGTCGTGTA | CAAAAAGACT | CTGTTTGTAG | 240 |
| AATTCACGGA | TCACCTTTTC | AACATCGCTA | AGCCAAGGCC | ACCCTGGATG | GGTCTGCTAG | 300 |
| GTCCTACCAT | CCAGGCTGAG | GTTTATGATA | CAGTGGTCAT | TACACTTAAG | AACATGGCTT | 360 |
| CCCATCCTGT | CAGTCTTCAT | GCTGTGGTG | TATCCTACTG | GAAAGCTTCT | GAGGGAGCTG | 420 |
| AATATGATGA | TCAGACCAGT | CAAAGGGAGA | AAGAAGATGA | TAAAGTCTTC | CCTGGTGGAA | 480 |
| GCCATACATA | TGTCTGGCAG | GTCCTGAAAG | AGAATGGTCC | AATGGCCTCT | GACCCACTGT | 540 |
| GCCTTACCTA | CTCATATCTT | TCTCATGTGG | ACCTGGTAAA | AGACTTGAAT | TCAGGCCTCA | 600 |
| TTGGAGCCCT | ACTAGTATGT | AGAGAAGGGA | GTCTGGCCAA | GGAAAAGACA | CAGACCTTGC | 660 |
| ACAAATTTAT | ACTACTTTTT | GCTGTATTTG | ATGAAGGGAA | AAGTTGGCAC | TCAGAAACAA | 720 |
| AGAACTCCTT | GATGCAGGAT | AGGGATGCTG | CATCTGCTCG | GGCCTGGCCT | AAAATGCACA | 780 |
| CAGTCAATGG | TTATGTAAAC | AGGTCTCTGC | CAGGTCTGAT | TGGATGCCAC | AGGAAATCAG | 840 |
| TCTATTGGCA | TGTGATTGGA | ATGGGCACCA | CTCCTGAAAGT | GCACTCAATA | TTCTCTGAAG | 900 |
| GTCACACATT | TCTTGTGAGG | AACCATCGCC | AGGCGTCCCT | GGAAATCTCG | CCAATAACTT | 960 |
| TCCTTACTGC | TCAAACACTC | TTGATGGACC | TTGGACAGTT | TCTACTGTTT | TGTCATATCT | 1020 |
| CTTCCCACCA | ACATGATGGC | ATGGAAGCTT | ATGTCAAAGT | AGACAGCTGT | CCAGAGGAAC | 1080 |
| CCCAACTACG | AATGAAAAAT | AATGAAGAAG | CGGAAGACTA | TGATGATGAT | CTTACTGATT | 1140 |
| CTGAAATGGA | TGTGGTCAGG | TTTGATGATG | ACAACTCTCC | TTCTTTTATC | CAAATTCGCT | 1200 |
| CAGTTGCCAA | GAAGCATCCT | AAAACCTGGG | TACATTACAT | TGCTGCTGAA | GAGGAGGACT | 1260 |
| GGGACTATGC | TCCCTTAGTC | CTCGCCCCCG | ATGACAGAAG | TTATAAAAGT | CAATATTTGA | 1320 |
| ACAATGGCCC | TCAGCGGATT | GGTAGGAAGT | ACAAAAAAGT | CCGATTTATG | GCATACACAG | 1380 |
| ATGAAACCTT | TAAGACTCGT | GAAGCTATTC | AGCATGAATC | AGGAATCTTG | GGACCTTTAC | 1440 |
| TTTATGGGGA | AGTTGGAGAC | ACACTGTTGA | TTATATTTAA | GAATCAAGCA | AGCAGACCAT | 1500 |
| ATAACATCTA | CCCTCACGGA | ATCACTGATG | TCCGTCCTTT | GTATTCAAGG | AGATTACCAA | 1560 |
| AAGGTGTAAA | ACATTTGAAG | GATTTTCCAA | TTCTGCCAGG | AGAAATATTC | AAATATAAAT | 1620 |
| GGACAGTGAC | TGTAGAAGAT | GGGCCAACTA | AATCAGATCC | TCGGTGCCTG | ACCCGCTATT | 1680 |
| ACTCTAGTTT | CGTTAATATG | GAGAGAGATC | TAGCTTCAGG | ACTCATTGGC | CCTCTCCTCA | 1740 |
| TCTGTACAAA | AGAATCTGTA | GATCAAAGAG | GAAACCAGAT | AATGTCAGAC | AAGAGGAATG | 1800 |
| TTCATCCTGTT | TTCTGTATTT | GATGAGAACC | GAAGCTGGTA | CCTCACAGAG | AATATAACAAC | 1860 |
| GCTTTCTCCC | CAATCCAGCT | GGAGTGCAGC | TTGAGGATCC | AGAGTTCCAA | GCCTCCAACA | 1920 |
| TCATGCACAG | CATCAATGGC | TATGTTTTTG | ATAGTTTGCA | GTTGTCAATT | TGTTTGCATG | 1980 |
| AGGTGGCATA | CTGGTACATT | CTAAGCATTG | GAGCACAGAC | TGACTTCCTT | TCTGTCTTCT | 2040 |
| TCTCTGGATA | TACCTTCAAA | CACAAAATGG | TCTATGAAGA | CACACTCACC | CTATTCCCAT | 2100 |
| TCTCAGGAGA | AACTGTCTTC | ATGTCGATGG | AAAACCCAGG | TCTATGGATT | CTGGGGTGCC | 2160 |
| ACAACTCAGA | CTTTCGGAAC | AGAGGCATGA | CCGCCTTACT | GAAGGTTTCT | AGTTGTGACA | 2220 |

| | | | | | | |
|-------------|-------------|------------|------------|-------------|-------------|------|
| AGAACACTGG | TGATTATTAC | GAGGACAGTT | ATGAAGATAT | TTCAGCATAC | TTGCTGAGTA | 2280 |
| AAAACAATGC | CATTGAACCA | AGAAGCTTCT | CCCAGAATTC | AAGACACCCCT | AGCACTAGGC | 2340 |
| AAAAGCAATT | TAATGCCACC | CCACCAGTCT | TGAAACGCCA | TCAACGGGAA | ATAACTCGTA | 2400 |
| CTACTCTTCA | GTCAGATCAA | GAGGAAATTG | ACTATGATGA | TACCATATCA | GTTGAAATGA | 2460 |
| AGAAGGAAGA | TTTTGACATT | TATGATGAGG | ATGAAAATCA | GAGCCCCCGC | AGCTTTCAAA | 2520 |
| AGAAAACACG | ACACTATTTT | ATTGCTGCAG | TGGAGAGGCT | CTGGGATTAT | GGGATGAGTA | 2580 |
| GCTCCCCACA | TGTTCTAAGA | AACAGGGCTC | AGAGTGGCAG | TGTCCCTCAG | TTCAAGAAAG | 2640 |
| TTGTTTTTCCA | GGAAATTTACT | GATGGCTCCT | TTACTCAGCC | CTTATAACCGT | GGAGAACTAA | 2700 |
| ATGAACATTT | GGGACTCCTG | GGGCCATATA | TAAGAGCAGA | AGTTGAAGAT | AATATCATGG | 2760 |
| TAACTTTTCAG | AAATCAGGCC | TCTCGTCCCT | ATTCTTTCTA | TTCTAGCCTT | ATTTCTTATG | 2820 |
| AGGAAGATCA | GAGGCAAGGA | GCAGAACCTA | GAAAAAACTT | TGTCAAGCCT | AATGAAACCA | 2880 |
| AAACTTACTT | TTGGAAAGTG | CAACATCATA | TGGCACCCAC | TAAAGATGAG | TTTGACTGCA | 2940 |
| AAGCCTGGGC | TTATTTCTCT | GATGTTGACC | TGGAAAAAGA | TGTGCACTCA | GGCCTGATTG | 3000 |
| GACCCCTTCT | GGTCTGCCAC | ACTAACACAC | TGAACCCTGC | TCATGGGAGA | CAAGTGACAG | 3060 |
| TACAGGAATT | TGCTCTGTTT | TTCACCATCT | TTGATGAGAC | CAAAAGCTGG | TACTTCACTG | 3120 |
| AAAATATGGA | AAGAAACTGC | AGGGCTCCCT | GCAATATCCA | GATGGAAGAT | CCCACTTTTA | 3180 |
| AAGAGAATTA | TCGCTTCCAT | GCAATCAATG | GCTACATAAT | GGATACACTA | CCTGGCTTAG | 3240 |
| TAATGGCTCA | GGATCAAAGG | ATTGATGTTT | ATCTGCTCAG | CATGGGCAGC | AATGAAAACA | 3300 |
| TCCATTCTAT | TCATTTTCAGT | GGACATGTGT | TCACTGTACG | AAAAAAAGAG | GAGTATAAAA | 3360 |
| TGGCACTGTA | CAATCTCTAT | CCAGGTGTTT | TTGAGACAGT | GGAAATGTTA | CCATCCAAAG | 3420 |
| CTGGAATTTG | GCGGGTGGAA | TGCCTTATTG | GCGAGCATCT | ACATGCTGGG | ATGAGCACAC | 3480 |
| TTTTTCTGGT | GTACAGCAAT | AAGTGTGAGA | CTCCCTGGG | AATGGCTTCT | GGACACATTA | 3540 |
| GAGATTTTCA | GATTACAGCT | TCAGGACAAT | ATGGACAGTG | GGCCCCAAAG | CTGGCCAGAC | 3600 |
| TTCATTATTC | CGGATCAATC | AATGCCTGGA | GCACCAAGGA | GCCCTTTTCT | TGGATCAAGG | 3660 |
| TGGATCTGTT | GGCACCAATG | ATTATTCACG | GCATCAAGAC | CCAGGGTGCC | CGTCAGAAGT | 3720 |
| TCTCCAGCCT | CTACATCTCT | CAGTTTATCA | TCATGTATAG | TCTTGATGGG | AAGAAGTGGC | 3780 |
| AGACTTATCG | AGGAATTTCC | ACTGGAACCT | TAATGGTCTT | CTTTGGCAAT | GTGGATTTCAT | 3840 |
| CTGGGATAAA | ACACAATATT | TTTAACCTC | CAATTATTGC | TCGATACATC | CGTTTGCAAC | 3900 |
| CAACTCATT | TAGCATTCGC | AGCACTCTTC | GCATGGAGTT | GATGGGCTGT | GATTTAAATA | 3960 |
| GTTGCAGCAT | GCCATTGGGA | ATGGAGAGTA | AAGCAATATC | AGATGCACAG | ATTACTGCTT | 4020 |
| CATCCTACTT | TACCAATATG | TTTGCCACCT | GGTCTCCTTC | AAAAGCTCGA | CTTCACCTCC | 4080 |
| AAGGGAGGAG | TAATGCCTGG | AGACCTCAGG | TGAATAATCC | AAAAGAGTGG | CTGCAAGTGG | 4140 |
| ACTTCCAGAA | GACAATGAAA | GTCACAGGAG | TAACACTCA | GGGAGTAAAA | TCTCTGCTTA | 4200 |
| CCAGCATGTA | TGTGAAGGAG | TTCTCATCT | CCAGCATCA | AGATGGCCAT | CAGTGGACTC | 4260 |
| TCTTTTTTCA | GAAATGGCAAA | GTAAAGGTTT | TTCAGGGAAA | TCAAGACTCC | TTCACACCTG | 4320 |
| TGGTGAACCT | TCTAGACCCA | CCGTTACTGA | CTCGCTACCT | TCGAATTAC | CCCCAGAGTT | 4380 |
| GGGTGCACCA | GATTGCCCTG | AGGATGGAGG | TTCTGGGCTG | CGAGGCACAG | GACCTCTACT | 4440 |
| GAGGGTGGCC | ACTGCAGCAC | CTGCCACTGC | CGTCACCTCT | CCCTCCTCAG | CTCCAGGGCA | 4500 |
| GTGTCCCTCC | CTGGCTTGCC | TTCTACCTTT | GTGCTAAATC | CTAGCAGACA | CTGCCCTTGAA | 4560 |
| GCCTCCTGAA | TTAACTATCA | TCAGTCCTGC | ATTTCTTTGG | TGGGGGGCCA | GGAGGGTGCA | 4620 |
| TCCAATTTAA | CTTAACTCTT | ACCGTCGACC | TGCAGGCCCA | ACGCGGCCGC | | 4670 |

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4451 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

| | | | | | | |
|------------|------------|------------|------------|-------------|------------|-----|
| AAGCTTAAAC | CATGCCCATG | GGGTCTCTGC | AACCGCTGGC | CACCTTGTAC | CTGCTGGGGA | 60 |
| TGCTGGTTCG | TTCCGTGCTA | GCCGCCACCC | GCCGCTACTA | CCTGGGCGCC | GTGGAGCTGT | 120 |
| CCTGGGACTA | CATGCAGAGC | GACCTGGGCG | AGCTCCCCGT | GGACGCCCGC | TTCCCCCCCC | 180 |
| GCGTGCCCAA | GAGCTTCCCC | TTCAACACCA | GCGTGGTGTA | CAAGAAAAACC | CTGTTCTGTG | 240 |
| AGTTACACGA | CCACCTGTTC | AACATTGCCA | AGCCGCGCCC | CCCCTGGATG | GGCCTGCTGG | 300 |
| GCCCCACCAT | CCAGGCCGAG | GTGTACGACA | CCGTGGTGAT | CACCCTGAAG | AACATGGCCA | 360 |
| GCCACCCCGT | GACCTGCAC | GCCGTGGGCG | TGAGCTACTG | GAAGGCCAGC | GAGGGCGCCG | 420 |
| AGTACGACGA | CCAGACGTCC | CAGCGCGAGA | AGGAGGACGA | CAAGGTGTTC | CCGGGGGGGA | 480 |
| GCCACACCTA | CGTGTGGCAG | GTGCTTAAGG | AGAACGGCCC | TATGGCCAGC | GACCCCTGT | 540 |
| GCCTGACCTA | CAGCTACCTG | AGCCACGTGG | ACCTGGTGAA | GGATCTGAAC | AGCGGGCTGA | 600 |
| TCGGCGCCCT | GCTGGTGTGT | CGCGAGGGCA | GCCTGGCCAA | GGAGAAAACC | CAGACCCTGC | 660 |

| | | | | | | |
|-------------|-------------|------------|------------|------------|-------------|------|
| ACAAGTTCAT | CCTGCTGTTC | GCCGTGTTCG | ACGAGGGGAA | GAGCTGGCAC | AGCGAGACTA | 720 |
| AGAACAGCCT | GATGCAGGAC | CGCGACGCCG | CCAGCGCCCC | CGCCTGGCCC | AAGATGCACA | 780 |
| CCGTAAACGG | CTACGTGAAC | CGCAGCCTGC | CCGGCCTGAT | CGGCTGCCAC | CGCAAGAGCG | 840 |
| TGTACTGGCA | CGTCATCGGC | ATGGGCACCA | CCCCTGAGGT | GCACAGCATC | TTCCTGGAGG | 900 |
| GCCACACCTT | CCTGGTGC GC | AACCACCGCC | AGGCCAGCCT | GGAGATCAGC | CCCATCACCT | 960 |
| TCCTGACTGC | CCAGACCTTG | CTGATGGACC | TAGGCCAGTT | CCTGCTGTTC | TGCCACATCA | 1020 |
| GCAGCCACCA | GCACGACGGC | ATGGAGGCTT | ACGTGAAGGT | GGACAGCTGC | CCCAGGAGC | 1080 |
| CCCAGCTGCG | CATGAAGAAC | AACGAGGAGG | CCGAGGACTA | CGACGACGAC | CTGACCGACA | 1140 |
| GCGAGATGGA | TGTCGTACGC | TTCGACGACG | ACAACAGCCC | CAGCTTCATC | CAGATCCGCA | 1200 |
| GCGTGGCCAA | GAAGCACCTT | AAGACCTGGG | TGCACTACAT | CGCCGCCGAG | GAGGAGGACT | 1260 |
| GGGACTACGC | CCCGCTAGTA | CTGGCCCCCG | ACGACCGCAG | CTACAAGAGC | CAGTACCTGA | 1320 |
| ACAACGGCCC | CCAGCGCATC | GGCCGCAAGT | ACAAGAAGGT | GCGCTTCATG | GCCTACACCG | 1380 |
| ACGAGACTTT | CAAGACCCGC | GAGGCCATCC | AGCACGAGTC | CGGCATCCTC | GGCCCCCTGC | 1440 |
| TGTACGGCGA | GGTGGGCGAC | ACCCTGCTGA | TCATCTTCAA | GAACCAGGCC | AGCAGGCCCT | 1500 |
| ACAACATCTA | CCCCACGGC | ATCACCGACG | TGCGCCCCCT | GTACAGCCGC | CGCCTGCCCA | 1560 |
| AGGGCGTGAA | GCACCTGAAG | GACTTCCCCA | TCCTGCCCGG | CGAGATCTTC | AAGTACAAGT | 1620 |
| GGACCGTGAC | CGTGAGGAGC | GGCCCCACCA | AGAGCGACCC | CCGCTGCCTG | ACCCGCTACT | 1680 |
| ACAGCAGCTT | CGTGAACATG | GAGCGCGACC | TGGCCTCCGG | ACTGATCGGC | CCCCTGCTGA | 1740 |
| TCTGCTACAA | GGAGAGCGTG | GACCAGCGCG | GCAACCAGAT | CATGAGCGAC | AAGCGCAACG | 1800 |
| TGATCCTGTT | CAGCGTGTTC | GACGAGAACC | GCAGCTGGTA | TCTGACCGAG | AACATCCAGC | 1860 |
| GCTTTCCTGCC | CAACCCCGCT | GGCGTGCAGC | TGGAAGATCC | CGAGTTCCAG | GCCAGCAACA | 1920 |
| TCATGACAGC | CATCAACGGC | TACGTGTTCG | ACAGCGTGCA | GCTGAGCGTG | TGCCTGCATG | 1980 |
| AGGTGGCCTA | CTGGTACATC | CTGAGCATCG | CGCGCCAGAC | CGACTTCCTG | AGCGTGTTCCT | 2040 |
| TCTCCGGGTA | TACCTTCAAG | CACAAGATGG | TGTACGAGGA | CACCTTGACC | CTGTTCCCCCT | 2100 |
| TCTCCGGCGA | GACTGTGTTC | ATGTCTATGG | AGAACCCCGG | CCTGTGGATT | CTGGGCTGCC | 2160 |
| ACAACAGCGA | CTTCCGCAAC | CGCGGCATGA | CTGCCCTGCT | GAAAGTCTCC | AGCTGCGACA | 2220 |
| AGAACACCGG | CGACTACTAC | GAGGACAGCT | ACGAGGACAT | CTCCGCCTAC | CTGCTGTCCA | 2280 |
| AGAACAACCG | CAATCAGCCC | CGCTCCTTCT | CCCAAAACTT | CCGCCACCCC | AGCACGCGTC | 2340 |
| AGAAGCAGTT | CATACGCCACC | CCCCCGTGC | TCCGCGCCGA | CCAGCCCGAG | ATCACCCGCA | 2400 |
| CCACCCTGCA | AAGCGACCAG | GAGGAGATCG | ACTACGACGA | CACCATCAGC | GTGGAGATGA | 2460 |
| AGAAGGAGGA | CTTCGACATC | TACGACGAGG | ACGAGAACCA | GAGCCCCCGC | TCCTTCCAAA | 2520 |
| AGAAAACCCG | CCACTACTTC | ATCGCCGCCG | TGGAGCGCCT | GTGGGACTAC | GGCATGAGCA | 2580 |
| GCAGCCCCCA | CGTCCTGCGC | AACCGCGCCC | AGAGCGGCAG | CGTGCCCCAG | TTCAAGAAGG | 2640 |
| TGGTGTTCCT | GGAGTTTACC | GACGGCAGCT | TCACCCAGCC | CCTGTACCGC | GGCGAGCTGA | 2700 |
| ACGAGCACCT | GGGCCTGCTC | GGCCCCTACA | TCCGCGCCGA | GGTGGAGGAC | AACATCATGG | 2760 |
| TGACCTTCCG | CAACCAAGCC | TCCCGGCCCT | ACTCCTTCTA | CTCCTCCCTG | ATCAGCTACG | 2820 |
| AGGAGGACCA | GCGCCAGGGC | GCCGAGCCCC | GCAAGAACTT | CGTGAAGCCC | AACGAGACTA | 2880 |
| AGACCTACTT | CTGGAAGGTG | CAGCACCACA | TGGCCCCCAC | CAAGGACGAG | TTCGACTGCA | 2940 |
| AGGCCTGGGC | CTACTTCAGC | GACGTGGACC | TGGAGAAGGA | CGTGCACAGC | GGCCTGATCG | 3000 |
| GCCCCCTGCT | GGTGTGCCAC | ACCAACACCC | TGAACCCCCC | CCACGGGAGG | CAGGTGACTG | 3060 |
| TGCAGGAATT | TGCCCTGTTC | TTCACCATCT | TCGACGAGAC | TAAGAGCTGG | TACTTCACCG | 3120 |
| AGAACATGGA | GCGCAACTGC | CGCGCCCCCT | GCAACATCCA | GATGGAAGAT | CCCACTTCA | 3180 |
| AGGAGAACTA | CCGCTTCCAC | GCCATCAACG | GCTACATCAT | GGACACCCTG | CCCGGCCTGG | 3240 |
| TGATGGCCCA | GGACCAGCGC | ATCCGCTGGT | ACCTGCTGTC | TATGGGCAGC | AACGAGAACA | 3300 |
| TCCACAGCAT | CCACTTCAGC | GGCCACGTTT | TCACCGTGCG | CAAGAAGGAG | GAGTACAAGA | 3360 |
| TGGCCCTGTA | CAACCTGTAC | CCCGGCGTGT | TCGAGACTGT | GGAGATGCTG | CCCAGCAAGG | 3420 |
| CCGGATCTG | GCGCGTGGAG | TGCCCTGATC | GCGAGCCTT | GCACGCCGGC | ATGAGCACCC | 3480 |
| TGTTCTTGGT | GTACAGCAAC | AAGTGCCAGA | CCCCCTGGG | CATGGCCAGC | GGCCACATCC | 3540 |
| GCGACTTCCA | GATCACCGCC | AGCGGCCAGT | ACGGCCAGTG | GGCTCCCAAG | CTGGCCCCGC | 3600 |
| TGCACTACAG | CGGCAGCATC | AACGCCTGGT | CGACCAAGGA | GCCCTTCTCC | TGGATCAAGG | 3660 |
| TGGACCTGCT | GGCCCCCATG | ATCATCCACG | GCATCAAGAC | CCAGGGCGCC | CGCCAGAAGT | 3720 |
| TCAGAGCCT | GTACATCAGC | CAGTTCATCA | TCATGTACTC | TCTAGACGGC | AAGAAGTGGC | 3780 |
| AGACCTACCG | CGGCAACAGC | ACCGGCACCC | TGATGGTGT | CTTCGGCAAC | GTGGACAGCA | 3840 |
| GCGGATCTAA | GCACAACATC | TTCAAACCCC | CCATCATCGC | CCGCTACATC | CCCTGCGACC | 3900 |
| CCACCCACTA | CAGCATCCGC | AGCACCTTGC | GCATGGAGCT | GATGGGCTGC | GACCTGAACA | 3960 |
| GCTGCAGCAT | GCCCCCTGGG | ATGGAGAGCA | AGGCCATCAG | CGACGCCCCG | ATCACCGCCT | 4020 |
| CCAGCTACTT | CACCAACATG | TTCGCCACCT | GGAGCCCCAG | CAAGGCCCGC | CTGCACCTGC | 4080 |
| AGGGCCGCGA | CAACGCCTGG | CGCCCCCAGG | TGAACAACCC | CAAGGAGTGG | CTGCAGGTGG | 4140 |
| ACTTCCAGAA | AACCTGAAG | GTGACTGGCG | TGACCACCCA | GGGCGTCAAG | AGCCTGCTGA | 4200 |
| CCAGCATGTA | CGTGAAGGAG | TTCCTGATCA | GCAGCGCCA | GGACGGCCAC | CAGTTGCACCC | 4260 |
| TGTTCTTCCA | AAACGGCAAG | GTGAAGGTGT | TCCAGGGCAA | CCAGGACAGC | TTACACCCGG | 4320 |
| TCGTGAACAG | CCTGGACCCC | CCCCTGCTGA | CCCGCTACCT | GCGCATCCAC | CCCCAGAGCT | 4380 |
| GGGTGCACCA | GATCGCCCTG | CGCATGGAGG | TGCTGGGCTG | CGAGGCCCCG | GACCTGTACT | 4440 |
| GAAGCGGCCG | C | | | | | 4451 |

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGGGATCCTC ACGTCTCA

18

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CTGCTTCTGA CGCGTGCTGG GGTGGCGGGA GTT

33

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

CTGCTGAAAG TCTCCAGCTG C

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

GGCAGGTGCT TAAGGAGAAC GGCCCTATGG CCA

33

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other